AD_____

Award Number: W81XWH-06-1-0131

TITLE: High-Throughput Analysis of Dynamic Gene Expression Associated with Sleep Deprivation and Recovery Sleep in the Mouse Brain

PRINCIPAL INVESTIGATOR: Ed Lein, Ph.D.

CONTRACTING ORGANIZATION: Allen Institute for Brain Science
Seattle, WA 98103

REPORT DATE: December 2006

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
01-12-2006	Annual	1 Dec 2005 – 30 Nov 2006

4. TITLE AND SUBTITLE

High-Throughput Analysis of Dynamic Gene Expression Associated with Sleep Deprivation and Recovery Sleep in the Mouse Brain

5a. CONTRACT NUMBER

5b. GRANT NUMBER
W81XWH-06-1-0131

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Ed Lein, Ph.D.

Email: edl@alleninstitute.org

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Allen Institute for Brain Science
Seattle, WA 98103

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES
Original contains colored plates: ALL DTIC reproductions will be in black and white.

14. ABSTRACT

Sleep deprivation in mice causes changes in gene expression in discrete brain regions associated with sleep/wake control as well as higher cognitive functions. Laser capture microdissection was used to isolate a set of relevant brain regions, and these samples were amplified to allow genome-wide genetic profiling on DNA microarrays to search for altered gene expression correlated with sleeping, waking, sleep deprivation and recovery sleep following sleep deprivation. Candidate genes from this microarray approach have been comprehensively analyzed using a high-throughput in situ hybridization platform. The resulting data indicate that there is dynamic gene expression correlated with circadian time of day, sleeping vs. waking states, and sleep deprivation specifically. This dynamic gene expression is also highly brain region and cell type-specific within a given region. These genes help to provide an anatomical and genetic framework for understanding the consequences of sleep deprivation and potential new means for pharmacological intervention. A web-based viewer interface has been developed to provide access to this data for the entire sleep community.

15. SUBJECT TERMS
Sleep deprivation, Recovery sleep, In situ hybridization, DNA microarray, Laser Capture Microdissection

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			USAMRMC
U	U	U	UU	33	19b. TELEPHONE NUMBER *(include area code)*

Table of Contents

Page

Introduction

Sleep deprivation leads to a repertoire of cognitive, attentional and emotional deficits that are seriously detrimental in professions requiring alertness (*1, 2*). While a great deal is known about the behavioral consequences of sleep deprivation, relatively little is known about the anatomical and molecular basis of the consequences of sleep deprivation within the brain. The robust nature of behavioral manipulations producing sleep deprivation in rodent model systems, coupled with the anatomical specificity of brain circuitry involved with sleep regulation and sleep deprivation-induced behavioral deficits, present an ideal biological paradigm for which the application of large-scale molecular and histological "discovery science" techniques can lend tremendous value. The goal of this study is to examine the molecular consequences of sleep deprivation and the restorative actions of sleep, with special regard to behavioral deficits associated with sleep deprivation and to mechanisms underlying sleep regulation. The approach involves three phases. First, 48 candidate genes originally described in the literature as responding to sleep deprivation or involved in sleep regulation were characterized by in situ hybridization after sleep deprivation and/or recovery sleep. Second, a combination of laser capture microdissection and DNA microarray is used to produce a systematic, genome-wide analysis of changes in gene expression in discrete brain nuclei that occur in response to sleep deprivation and recovery sleep following sleep deprivation. Lastly, new candidate genes showing altered gene expression as a function of behavioral state and/or brain region are then comprehensively analyzed across the entire brain and across behavioral conditions using the high-throughput *in situ* hybridization platform generated by the Allen Brain Atlas, a genome-scale atlas of gene expression in the mouse brain. In total, these data should provide functionally relevant, anatomically-specific data addressing the molecular and anatomic architecture underlying sleeping and waking states, the deficits associated with sleep deprivation and the restorative effects of sleep. Data from the entire *in situ* hybridization component of this study will be made available to the sleep community on a publicly accessible web site (www.brainatlas.org).

Body

The large-scale analysis of sleep deprivation- and recovery sleep-associated gene expression changes in the current project is divided into two main aims as detailed in the original Statement of Work. Broadly, the first Aim, now completed, involved the development and validation of a behavioral paradigm for producing sleep deprivation in C57BL/6 mice, *in situ* hybridization analysis of genes derived from the sleep literature across this behavioral paradigm, and the development of tools to allow analysis of dynamic gene expression across the resulting large data set. The second Aim involves a phase of technique development to allow laser capture microdissection of discrete brain regions and subsequent amplification methodologies for small RNA sample inputs for microarray applications (complete). This methodology is applied to genetically profile a series of sleep/wake or sleep-deprivation associated brain regions across the behavioral paradigm from Aim 1, allowing a systematic search for genes that show behavioral state and/or brain region-specific changes in gene expression (largely complete). Finally, a large cohort of these candidate genes are processed for *in situ* hybridization to generate the bulk of the data set for the project (in progress). Qualitative and quantitative analysis of this cellular resolution ISH data is occurring in parallel with data production, and leads to confirmation of microarray results and subsequent reprocessing to generate replicate data sets. Data from the first 48 genes processed is now available for public viewing through the Allen Brain Atlas web site (www.brainatlas.org). This annual report is structured in the format of the original Statement of Work to detail progress on each defined specific aim.

Aim 1: Sleep/wake and sleep deprivation methodology development, and *in situ* hybridization of candidate genes derived from the literature (Months 1-6)

Aim 1.1: Validation of a controlled system for producing sleep deprivation.

Mice display a pattern of sleep activity that is much more evenly distributed across the entire 24-hour cycle than humans (*3*). Sleep states are typically defined on the basis of electromyographic (EMG) and electroencephalographic (EEG) signatures, which allow the identification of rapid eye movement (REM) and non-rapid eye movement sleep (NREM). In mice, the bulk of both REM and NREM sleep occur during the light phase of the 24-hour cycle (*3*). The C57BL/6J strain used in this study exhibits a second sleep-predominant phase late in the second half of the dark phase of the 24-hour cycle (*4, 5*) (**Figure 1A**).

Numerous paradigms have been utilized to prevent animals from sleeping for sleep deprivation studies. Generally these involve some form of non-stressful, gentle manipulation of the animal's cage or physical handling of the animal directly, either following periods of inactivity or when online EEG recording indicates that the animal has begun to sleep (*5-8*). The timing and length of sleep deprivation were important factors to consider given the goals of the study. The major sleep period for mice begins soon after lights on (*7*). Therefore, to achieve the largest sleep deficit, sleep deprivation should begin at lights on, or ZT0. There is evidence that sleep deprivation periods of 3 hrs generally result in changes in immediate-early gene expression (*9*), while longer periods of sleep deprivation result in altered regulation of many other classes of transcripts, including vesicle- and synapse-related genes, receptors, transporters, and enzymes (*9*). Therefore, to study the effects of sleep deprivation on sleep regulatory centers as well as on genes associated with cognitive, attentional, and emotional deficits, longer sleep deprivation periods are considered optimal, and in mice, 6-8 hours is the practical limit for sleep deprivation largely due to the fact that it is technically challenging to keep mice awake for longer than this (*6*).

Based on these considerations, five behavioral conditions were selected for the current study (**Figure 1B**). All experimental animal production and behavioral analysis was performed as a subcontract in the laboratory of Dr. Thomas Kilduff at SRI International, using methods routinely performed in the Kilduff laboratory at SRI International to study the effects of sleep deprivation on gene expression in the brain in rats (*5, 10, 11*). New implants for EEG/EMG recording in mice were tested and optimized for this study, by laboratory staff with prior experience in sleep recording in mice. Mice begin sleep deprivation by gentle handling at ZT0, and are sacrificed at the completion of the 6 hr sleep deprivation. Appropriate cage controls are also sacrificed at this

time (ZT6). Some mice are permitted recovery sleep of 4 hours ending at ZT10; this length of time should enable the detection of several types of affected genes in addition to the expected immediate-early genes. Appropriate cage controls are also sacrificed at ZT10 to enable dissociation of dynamic circadian gene expression from sleep deprivation- or recovery sleep-induced genes. One additional group of mice is sacrificed during the night phase to obtain mice which are spontaneously awake. This time point serves not only as a control for genes associated with waking, but also as a control for genes that oscillate with circadian rhythmicity.

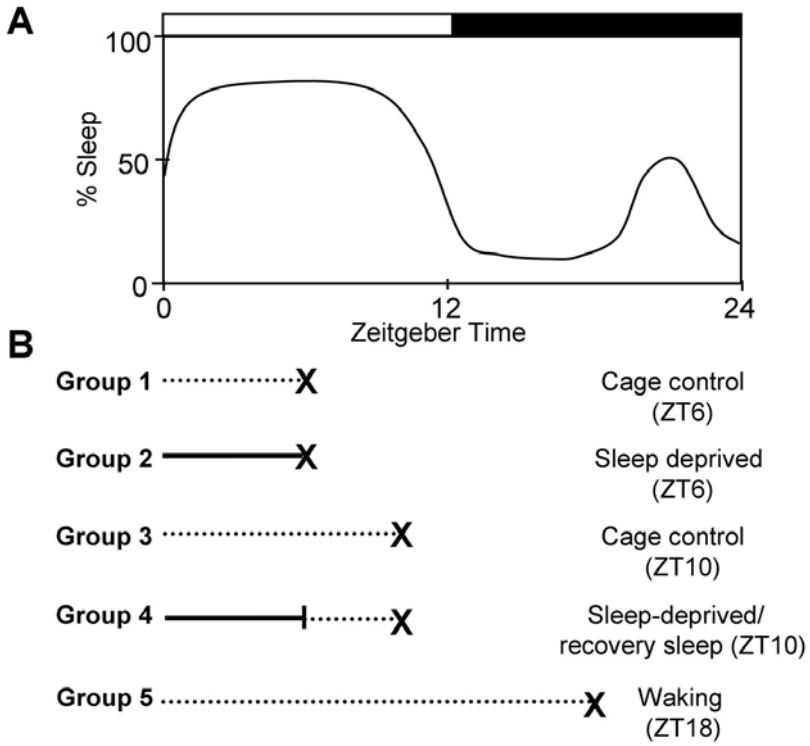


Figure 1. Experimental design (A) Schematic of sleeping and waking activity as a function of light/dark cycles in C57BL/6J mice [modified from Figure 1 of (*5*)]. (B) Experimental groups for sleep deprivation. Schematic is drawn with reference to the time scale as shown in (A). Dotted lines indicate time spent under standard housing conditions. Solid lines denote gentle handling treatment for sleep deprivation. Time of animal sacrifice is marked with an "X."

The initial set of experimental mice were implanted for EEG/EMG recording and sleep/wake data analysis to confirm normal baseline sleep/wake cycles as reported by other laboratories (*5*, *7*). Specifically, EEG/EMG recordings were used to validate that gentle handling techniques adequately prevent significant periods of sleep and to confirm that sleep deprivation results in intense recovery sleep. 50 male C57BL/6J mice approximately 2 months of age (10 mice per condition) were implanted for EEG/EMG recording. Data was recorded for both baseline and experimental days, and sleep as a percentage of time was calculated for these animals. As shown in **Figure 2**, the sleep deprivation paradigm effectively eliminated significant periods of sleep, while allowing animals to sleep following sleep deprivation led to a rapid rebound in the percentage of time asleep. Analysis of EEG power spectra, a better measure of sleep debt, demonstrates increased sleep need immediately following sleep deprivation (**Figure 3**).

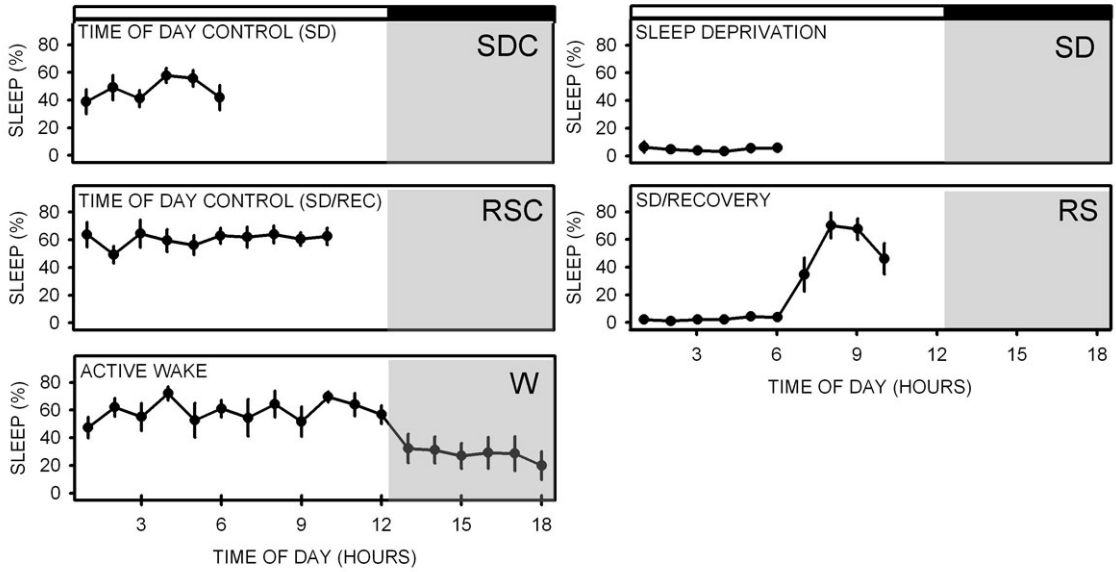


Figure 2. Percentage of sleep over time across experimental behavioral conditions.

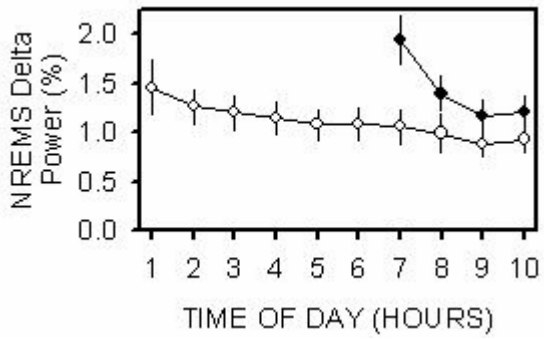


Figure 3. NREM delta power over time in normal animals (white circles) and animals allowed to sleep following six hours of sleep deprivation.

Detailed experimental protocols for producing sleep deprivation, generating and analyzing EEG/EMG data are provided below. These data were considered sufficient to validate the efficacy of the behavioral paradigm for the remainder of the project. Subsequently generated animals were treated identically but not implanted for recording. The original plan had been to use the initial implanted set of animals for laser capture microdissection and microarray analysis. However, removal of the EEG implants leads to significant damage to the underlying neocortex; since the neocortex was one of the structures targeted for microarray analysis, non-implanted animals were used for microarray analysis instead. The implanted brains were instead used for *in situ* hybridization experiments as described below.

Experimental Methodology

Experimental Animals. 8-week old adult male C57BL/6J mice (the same age, sex and strain used for the Allen Brain Atlas project to allow comparison) were purchased from Jackson Laboratory (Bar Harbor, ME). Animals were housed in polycarbonate cages and routinely entrained to a 12 h light/12 h dark cycle (LD 12:12) for at least 3 weeks before use in an experiment; they were given free access to food and water. At the end of the sleep deprivation (SD) and recovery sleep (RS) periods, experimental mice and equal numbers of control animals were sacrificed by decapitation. The brains were dissected, frozen on dry ice, and stored at -70°C. For behavioral validation of sleep deprivation protocols, 50 mice were implanted for EEG and EMG recordings, and of these, 11 mice were sacrificed in each of five conditions: sleep deprivation (ZT6), cage control (ZT6), recovery sleep (ZT10), cage control (ZT10), and waking (ZT16-18). After validation of the effects of sleep

deprivation on EEG slow wave activity in these mice, an additional 63 animals were sacrificed during each of the 5 experimental conditions without implantation for EEG/EMG recordings (see "Numbers of Animals" below).

EEG and EMG Surgical Implantation. Under Isoflurane anesthesia, mice were prepared with a chronic recording implant that permits continuous EEG and EMG recordings. Stainless steel screws (#000) affix the implant to the skull and serve as epidural electrodes (2 frontal and 2 occipital). The screws are attached to Teflon-coated stainless steel wires. Dental acrylic is used to further affix the implant to the skull. Multi-stranded, twisted stainless steel wire electrodes are sutured bilaterally in the neck muscles for recording the EMG. Incisions are closed in layers with absorbable suture (Vicryl 3-0) and antibiotics administered topically and systemically. Animals are permitted a 2-week postoperative recovery before study. During recordings, one of the four electrodes was used as an animal ground to reduce electrical noise.

Sleep Recordings and Analysis. Animals were connected via a cable and a counterbalanced commutator to an Embla data collection system (ResMed, Poway, CA). The amplified and digitized signals are stored and analyzed on a computer using Somnologica software (v. 2.0.1, ResMed, Poway, CA) or SleepSign (Kissei Comtec, Irvine CA). EEG and EMG data was scored visually for waking (W), rapid eye movement sleep (REMS), and non-rapid eye movement sleep (NREMS). Waking is defined by low voltage, low theta EEG with high voltage EMG. REM sleep is defined by the lowest voltage EMG and low voltage EEG with high theta activity. Non-REM sleep is all sleep other than REM sleep. Total sleep time was calculated, as well as the following parameters for each state: (1) hourly totals for the 6 hrs of SD (SD group); the 6 hrs corresponding to the SD session (ZT6 controls); the 6 hrs of SD (SD group) and 4 hrs of RS (RS group); the 10 hrs corresponding to SD and RS (ZT10 controls); and 18 hours beginning at ZT0 (ZT18 group).

Numbers of Animals. 50 mice were implanted for EEG and EMG recording (10 mice X 5 groups). These animals will later be used for LCM studies and/or ISH studies. An additional 330 mice (63 mice/condition x 5 conditions were sacrificed without surgical implantation.

SD Procedures. When EEG recordings and visual observations suggested mice are about to sleep, animals were kept awake by a combination of cage tapping, introduction of foreign objects (e.g., balled paper towels), cage rotation, and stroking of vibrissae and fur with a brush. At the end of the SD periods, experimental and control mice were sacrificed by cervical dislocation and decapitation. Such "gentle handling" procedures are a time-honored SD technique and have been used extensively by other sleep laboratories such as that of Drs. Borbely and Tobler of the Pharmacologisches Institut der Zurich (Switzerland) in studies of EEG delta power (*12-15*)

.

Aim 1.2: Molecular validation of gene expression changes produced by the sleep deprivation methodology.

The initial *in situ* hybridization (ISH) phase of the project was performed on a set of candidate genes described in the existing literature to change as a function of sleep or sleep deprivation. These genes were examined throughout the entire brain across the experimental conditions described in Aim 1.1. 48 genes culled from the literature (**Appendix I**) were processed on the brains that had been implanted for EEG/EMG recordings. These data were generated in a manner identical to that used for the Allen Brain Atlas project (Lein et al., 2006; see also www.brainatlas.org), using high-throughput semiautomated tissue processing. Non-isotopic digoxigenin-labeled probes for each gene were hybridized to 5 sagittal slides per brain (20 - 25 um sections spanning an entire hemisphere), allowing sampling across all major brain regions, for each of the 5 experimental conditions. This method uses alkaline phosphatase and NBT/BCIP to generate purple punctate labeling in cells expressing a particular transcript, that can be visualized using brightfield microscopy. 10x tiles are collected across

individual tissue sections using an automated image capture system. These tiles are subsequently stitched and compressed using JPEG2000 compression to produce high resolution montages of each section at a size amenable to web-based viewing.

Many of the genes identified through the literature can be classified as immediate early genes (IEGs). IEG activation is thought to mirror physiological activation, and thereby provides a readout of functional circuitry activated by a particular experimental condition. Therefore, this initial phase was essentially an IEG mapping study, identifying brain regions that are selectively activated by one or more of the experimental variables (circadian time, sleep/wake state, sleep deprivation and recovery sleep). **Figure 4** shows IEG expression patterns across the five experimental conditions. Even at low magnification there is clear behavioral state-dependent IEG activation. For example, expression in visual cortex is much lower in the waking condition than the other conditions, likely reflecting the absence of light-driven retinal activation selectively in this condition. Many other differences are seen, corresponding with, for example, waking (up in waking and sleep deprivation), and are often very discrete in their localization to small nuclei or subdivisions of nuclei (for example, the dorsal margin of the caudate putamen). Interestingly, different IEGs show somewhat different patterns. This may reflect the different timecourse of activation of different IEGs. Notably, not all candidate genes derived from the literature showed robust sleep/wake- or sleep-deprivation associated changes, and the IEGs were the most successful category among these candidates, including Arc, Egr1, Egr3, Fos, Fosb, Fosl2, Homer1, JunB, Nr4a1 and Nr4a3. There are many potential reasons for this discrepancy. First, many of these genes were microarray-derived with very mild fold change enrichment, and should be treated as candidates until validated with other methods. Second, many of the candidates were derived from rat data instead of mouse data, using different experimental paradigms. Third, some of these genes were reported on the basis of quantitative RT-PCR results, which is more sensitive to low level transcripts and to modest changes in expression than ISH. Finally, there may be a sensitivity threshold with the ISH methodology used in the current project, although this seems somewhat unlikely given the good concordance between Allen Brain Atlas data and other data sources (Lein et al., 2006).

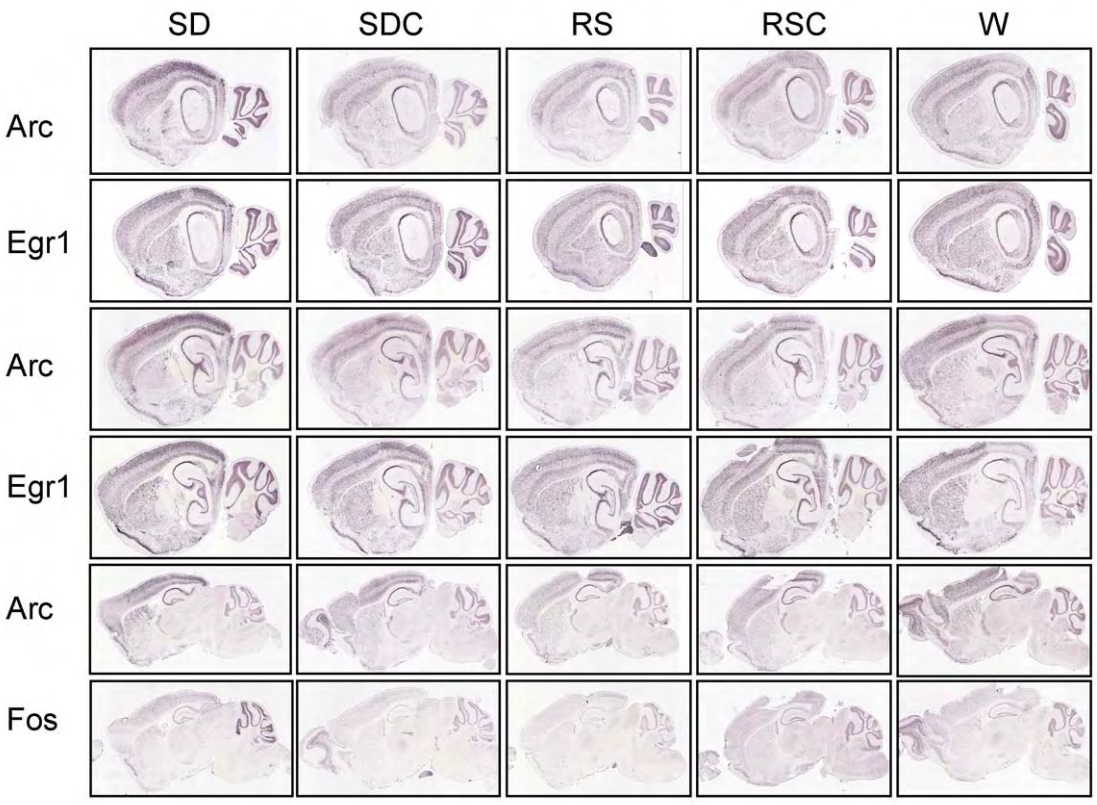


Figure 4. Expression of three immediate-early genes, Arc, Egr1 and Fos, across the five behavioral conditions. Gross differences in gene expression are visible across the brain. SD, sleep deprivation; SDC, sleep deprivation control; RS, recovery sleep; RSC, recovery sleep control; W, wake.

A major goal for this phase of the project was to identify very specific brain regions affected by sleep deprivation for subsequent microarray analysis. Specifically, we focused on regions of the neocortex, amygdala and hippocampus that showed differential activation between sleep deprivation and recovery sleep, since these structures correlate with the known detrimental effects of sleep deprivation on cognitive, emotional and mnemonic abilities (**Figure 5**). Consistent with this prediction, we found IEG upregulation specifically in orbitofrontal cortex. Within the amygdala, the central nucleus has been implicated in sleep deprivation; however, we did not observe changes in the central nucleus, but rather found robust changes in the neighboring amydalocortical nuclei. Finally, only modest changes in IEG expression were seen in the hippocampus proper (Ammon's horn and dentate gyrus), while there is robust upregulation in another part of the hippocampal formation, the entorhinal cortex that provides the major input to the dentate gyrus. These findings led to the selection of orbitofrontal cortex, the amygdalocortical nuclei, and the entorhinal cortex for subsequent laser capture and microarray analysis. The entire data set for the 48 literature-derived candidates is now publicly accessible through the Allen Brain Atlas web site (www.brainatlas.org).

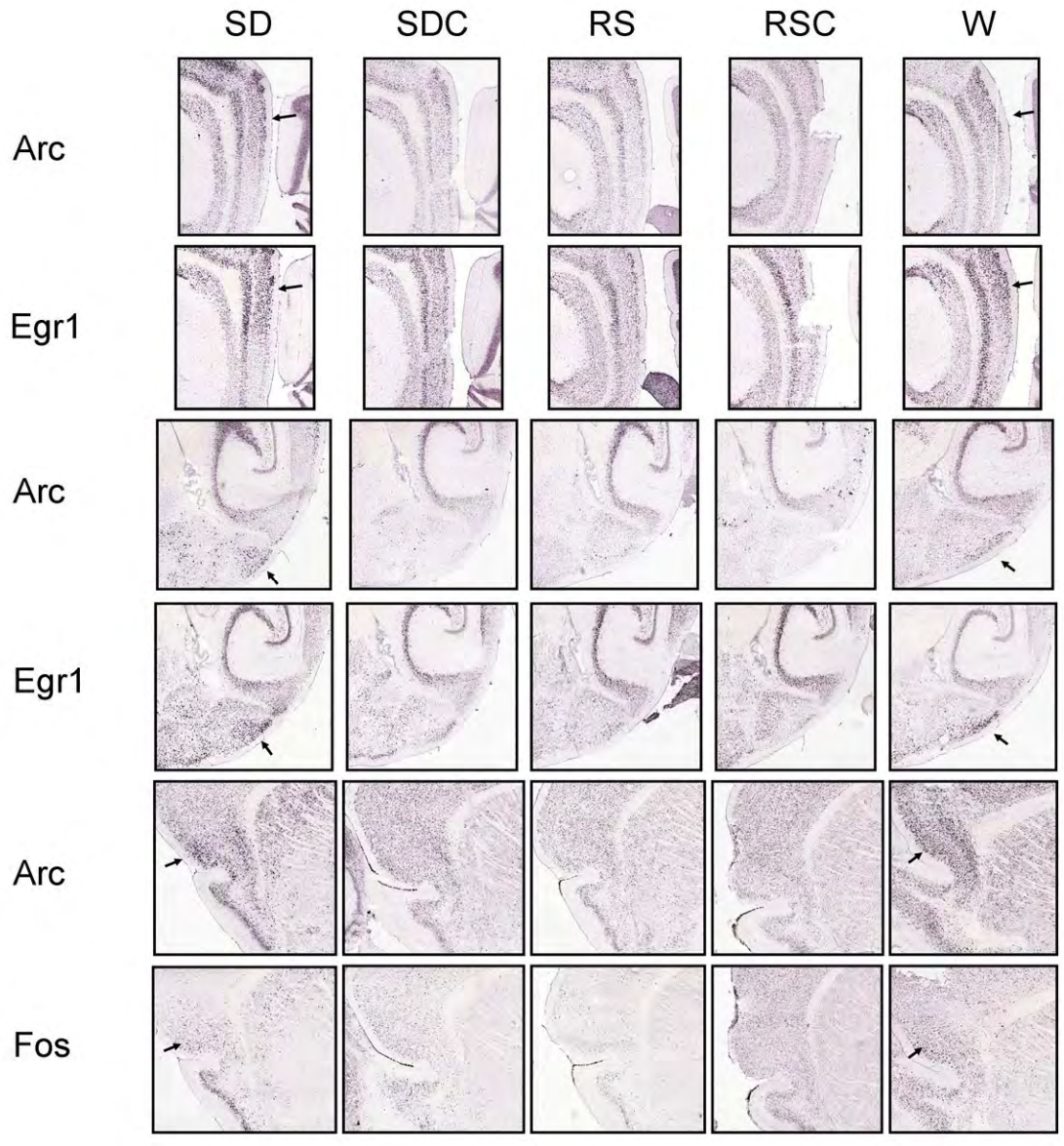


Figure 5. High magnification images of brain regions showing increased immediate early gene expression correlated with sleep deprivation. Induction of Arc and Egr1 in the dorsal part of the entorhinal cortex is associated with sleep deprivation and wake (arrows; first two rows). Induction of Arc and Egr1 in the posteromedial cortical amygdala are shown in SD and W; note differences in layer-specific expression within this structure between SD and W (arrows, 3rd and 4th rows). Induction of Arc and Fos in orbital cortex in sleep-deprivation and wake; note differences in signal distribution and intensity between SD and W (arrows, 5th and 6th rows).

Quantitative RT-PCR. A pilot study was completed to compare the expression of genes from ISH to values generated by quantitative RT-PCR from samples obtained by laser microdissection (**Figure 6**). Briefly, dorsal caudate putamen was collected from three of the experimental conditions: sleep-deprived, sleep-deprived cage control, and waking. The RNA was isolated and amplified for quantitative RT-PCR and visualized using SYBR green on the Roche Lightcycler. Generally these results corroborate the ISH data (e.g. Fos and Arc upregulated in sleep deprivation), although more subtle changes do not always perfectly match between the two conditions.

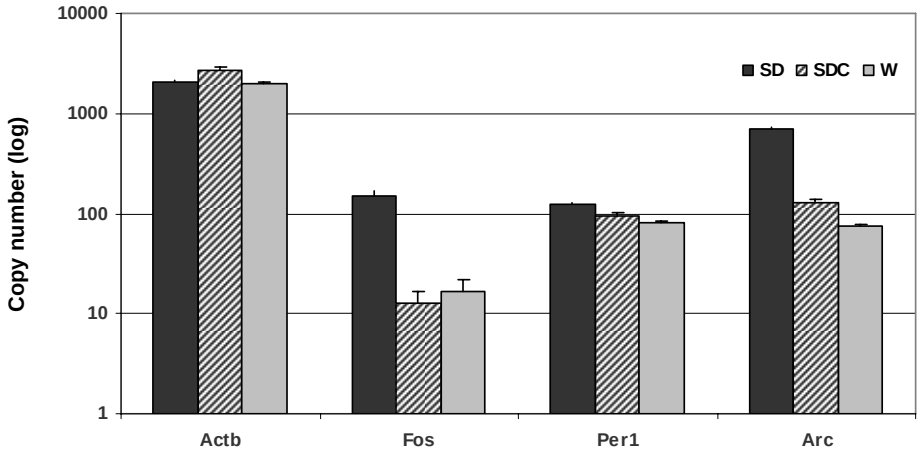


Figure 6. Copy number was calculated from quantitative RT-PCR for four genes. Each column is the average of 3 biological replicates, which were each assayed 4 times. ISH data showed upregulation of Fos and Arc in the sleep deprived (SD) condition, but not for Per1. Beta actin (Actb) is included as a control gene that is not expected to change across behavioral condition.

Aim 1.3: Development of improved statistical tools for registration and annotation of *in situ* hybridization data and for quantification of ISH signal across different behavioral conditions.

Large-scale high resolution histological data presents significant data analysis challenges. A suite of methods have been developed as part of the Allen Brain Atlas project that are being applied to these data, while new methods have been developed specifically for the current project. Briefly, the existing methodologies are based on registering ISH data to a 3D annotated digital model of the brain produced from an age- and strain-matched specimen. Automated Informatics methodologies are used to identify and quantify ISH signal on individual sections, and to register the series of images from a brain on to the reference model. A pseudocolored "heat mask" representing cellular expression for each section is produced in this process, and is also available on the web presentation of these data. The consequence of the registration to the reference model is that gene expression can be quantified for a series of 3D voxels (or summed across defined structures) spanning the entire brain for each gene in each experimental condition. 3D representations of gene expression across the entire brain, using a custom application called Brain Explorer, are also available on the web site, and provide a much more global view of expression across the entire brain than is possible looking at individual section.

We have developed new tools to allow comparison between samples as part of the current project. The registration of all data to a common anatomical framework described above allows a calculation of differential expression on a voxel-by-voxel basis across the entire brain. The resulting "difference masks," representing either absolute differences in expression/voxel or fold changes/voxel, can be plotted in 3D coordinates to give a rapid assessment of altered gene expression in discrete brain regions. An example of this type of visualization is shown in **Figure 7**, representing differential c-fos expression between sleep-deprived and the circadian time-matched control. One difficulty with this strategy on the initial set of brains is that the cortical damage noted earlier causes significant registration problems to the reference atlas model. Many of the IEGs have now been processed for replicate data sets on undamaged brains, with much better registration results.

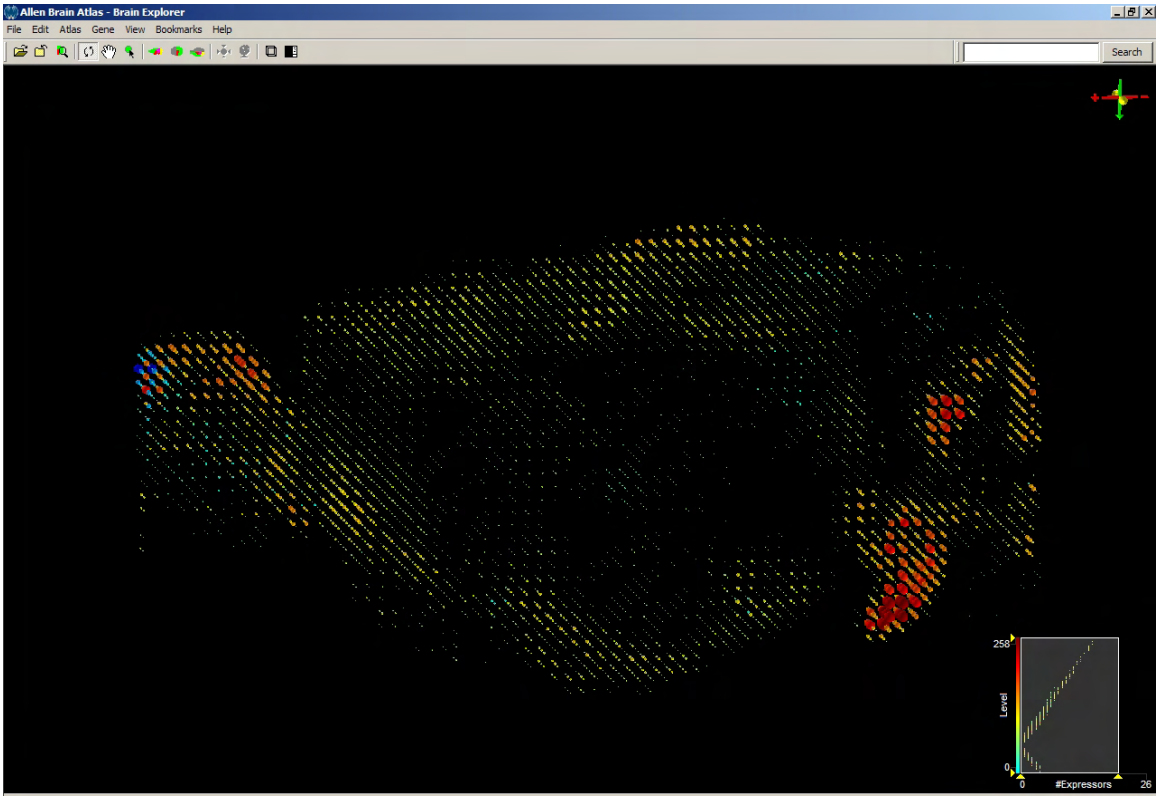


Figure 7. A screenshot of a difference mask for Fos from sleep-deprived (SD) brain and Fos from ZT6 cage control. Yellow and red indicate upregulation of signal intensity in SD condition. Blue represents downregulation of Fos signal intensity in SD condition. The size of the voxels indicate change in fractional area of expressing cells; larger voxels indicate an increase, which may be due to increase in size or in number of expressing cells.

While the informatics-based approach above is useful for assessing global trends in the data, it is limited by the resolution of the voxels (100-300um/side) and the fidelity of the registration. More traditional optical density and cell count measurements using ImagePro have now been adapted to allow targeted analysis in specific structures, and efforts are underway to semi-automate this process following manual delineation of target structures on the ISH or Nissl image. Data analysis in the later phases of this project will use a combination of informatics based screening for brain regions demonstrating changes, and more rigorous optical density and cell count measurements in those regions of interest.

Aim 2: Analysis of gene expression across sleep/wake cycles and following sleep deprivation (Months 2-18)

Aim 2.1: Laser capture microdissection (LCM) and microarray analysis.

Genetic profiling of isolated brain regions from the same experimental animals used for behavioral platform validation above in Aim 1.1

Sleep/Wake State Control

- Sleep-promoting: Hypothalamic preoptic area VLPO
- Arousal-promoting: Locus coeruleus, tuberomammillary nucleus, orexinergic neurons of the tuberal hypothalamus
- Circadian control: Suprachiasmatic nucleus

Sleep Deprivation Deficits

- Learning and memory: Hippocampus
- Emotional behavior: Amygdala
- Attentional, cognitive behaviors: Neocortex

The major part of this study involves a discovery science approach to genetically profile regions of the brain that are involved in sleep/wake regulation, circadian rhythmicity, and deficits associated with sleep deprivations. While these latter deficits are associated with large, heterogeneous structures including the hippocampus, amygdala and neocortex, the precise subdivisions to be analyzed were to be determined based on the behavioral state-dependent upregulation of immediate early genes described above. As predicted, the immediate early gene mapping indicated that alterations in gene expression occur in highly discrete subnuclei of these structures, although these regions differ from those previously described in the literature. For example, the input region for the hippocampus, showed much more dramatic activation than the hippocampus proper. The corticoamygdalar nuclei showed increased expression, and the frontal (orbitofrontal) cortex showed selective cortical activation with sleep deprivation, consistent with the role of frontal cortex in higher executive brain function. The final list of structures isolated for microarray analysis is shown in Table 1. Tissue from an additional set of structures implicated in sleep/wake control or sleep deprivation, or showing particular immediate early gene activation, were also isolated for possible quantitative PCR analysis at the end of the project.

Table 1. Regions for laser microdissection

Region	Abbreviation
Locus coeruleus	LC
Suprachiasmatic nucleus	SCN
Ventrolateral preoptic	VLPO
Hypocretin area	HCRT
Tuberomammillary nucleus	TMN
Posteromedial cortical amygdala	PMCO
Dorsal medial entrohinal cortex	MENT
Orbital cortex	ORB
Additional regions:	
Dorsal dentate gyrus	dDG
Dorsal CA3	dCA3
CA1	CA1
Central nucleus of amygdala	CEA

Basolateral amygdala	BLA
Laterodorsal tegmental area	LDTg
Basal forebrain	MCPO, SI, HDB

A significant amount of method development was necessary to allow laser capture microdissection/microarray analysis of the above brain regions across our behavioral paradigm. After extensive research the Leica LMD6000 laser microdissection system was selected as the optimal system for its contamination-free isolation, robust laser, powerful software and ease of use. This system uses a diode laser to cut through tissue mounted on film-coated slides, allowing isolation of single cells or discrete brain regions.

Tissue Preparation

Obtaining high quality, reproducible microarray data from small tissue samples using laser capture microdissection requires rapid processing of fresh cryostat tissue sections to allow sufficient morphology to identify structures of interest while maintaining high RNA quality. For practical purposes, it was necessary to be able to section tissue from experimental animals and freeze this tissue until the time of laser capture microdissection. Better results were obtained freezing the tissue immediately after sectioning, and then staining upon thawing. Section thickness proved to be critical, with decreasing morphological preservation at tissue thicknesses greater than 10µm due to the appearance of bubbles in the tissue, presumably due to formation of ice crystals (**Figure 8**). Cresyl violet staining is compatible with high quality RNA isolation. We tested several protocols, including a commercially available staining kit (Ambion). Comparable results were obtained using the staining kit and protocols developed previously; therefore, the earlier protocols were adopted, involving a rapid cold 70% ethanol fixation and subsequent 0.7% cresyl violet staining.

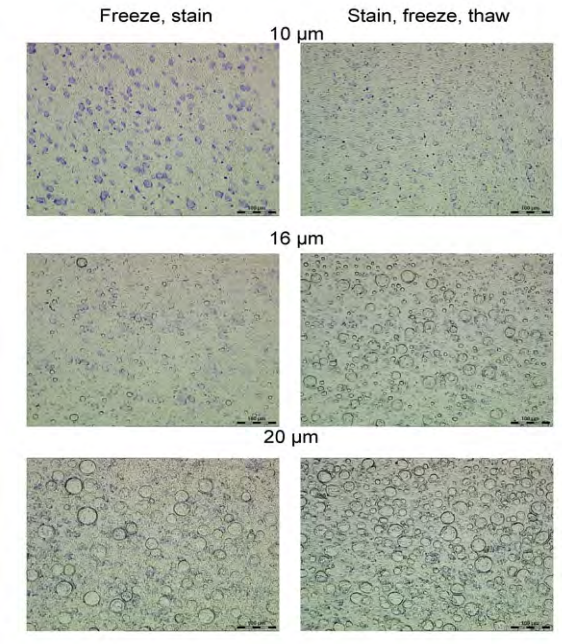


Figure 8. Comparison of section thicknesses (10, 16, 25 um) and slide handling methods (freezing and staining).

RNA Isolation

RNA isolation kits were compared from Qiagen, Arcturus, and Stratagene. The Qiagen RNeasy kit reproducibly produced high quality RNA following a DNase step to remove genomic DNA. Briefly, samples are microdissected into the Qiagen RLT buffer with fresh beta-mercaptoethanol added. Samples are capped, inverted and centrifuged, volume adjusted, vortexed for 30 seconds, centrifuged, and frozen on dry ice. RNA is isolated according to manufacturer's directions. RNA integrity and concentrated is evaluated using the Bioanalyzer Picochip with a 1 ng fresh RNA standard for normalization (**Figure 9**). RNA quality following laser microdissection is exceptionally good with this protocol as assessed by the integrity of 28S and 18S ribosomal RNA bands.

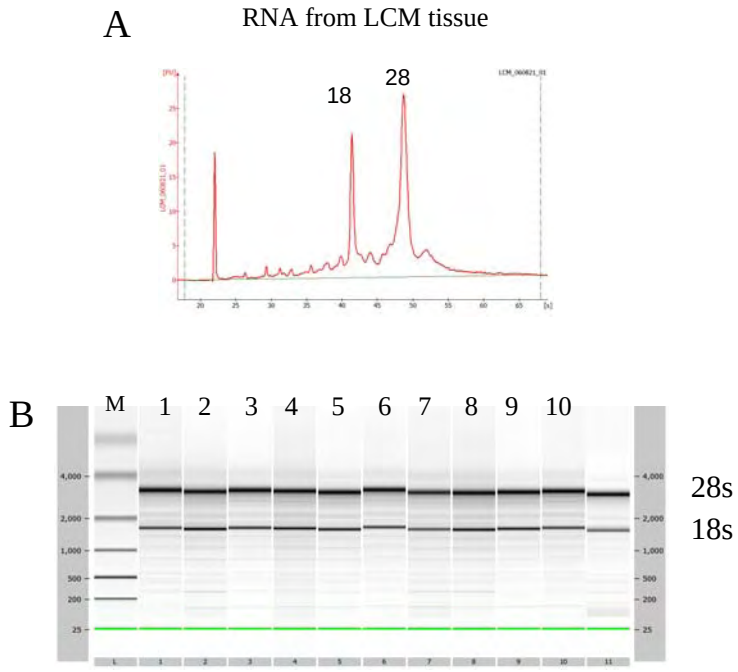


Figure 9. RNA quality after laser microdissection and isolation with Qiagen RNeasy. **A**, electropherogram of laser microdissected sample. **B**, simulated gel image from Bioanalyzer Picochip analysis. Lanes 1-10 are orbital cortex and entorhinal cortex laser microdissection samples. The last unmarked lane on the right is RNA isolated from fresh tissue for comparison.

Laser Capture Microdissection:

Isolation of the brain regions in Table 1 was achieved using laser capture microdissection (LCM) on tissues from experimental animals. The original goal was to achieve quadruplicate biological replicates from eight discrete brain regions across the five behavioral conditions. In most cases, histological appearance of these regions in cresyl violet-stained tissues was sufficient to reliably identify regions for isolation. However, the hypocretin-positive region of the tuberal hypothalamus is not easily identifiable by Nissl staining alone. To circumvent this problem, we performed *in situ* hybridization (ISH) for hypocretin (*Hcrt*) on adjacent sections to those for laser capture (1 ISH section per 3 LCM sections). A software feature of the Leica platform allows the transfer of a drawn region from one section to a serial, unstained section (**Figure 10**). All other areas were identified on the basis of cresyl violet staining alone.

Samples for the regions listed in Table 1 were isolated by laser microdissection, prepped using Qiagen RNeasy kit, and quantified using Bioanalyzer Picochips. Surprisingly, the RNA yields did not correspond to the amount of tissue collected, but corresponded instead to the sleep condition in a region specific manner (**Figure 11**). These data may indicate that ribosomal RNA, which is the major constituent of total RNA, may change with sleep condition in order to regulate the capacity for translation in the cell. This result has practical implications for this project: 1) it has been difficult to get sufficient RNA for amplification from certain conditions, such as the SCN sleep deprived control; 2) some of the lower samples fail out in amplification despite the existence of sufficient RNA (e.g., sleep-deprived entorhinal cortex). This issue is discussed in more detail below.

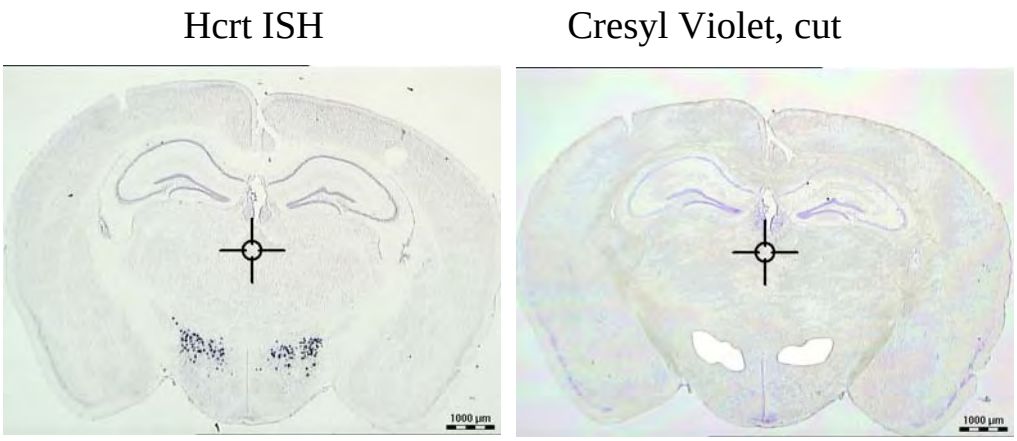


Figure 10. An ISH section probed for *Hcrt* was used to select regions for laser microdissection in an adjacent, cresyl violet-stained section.

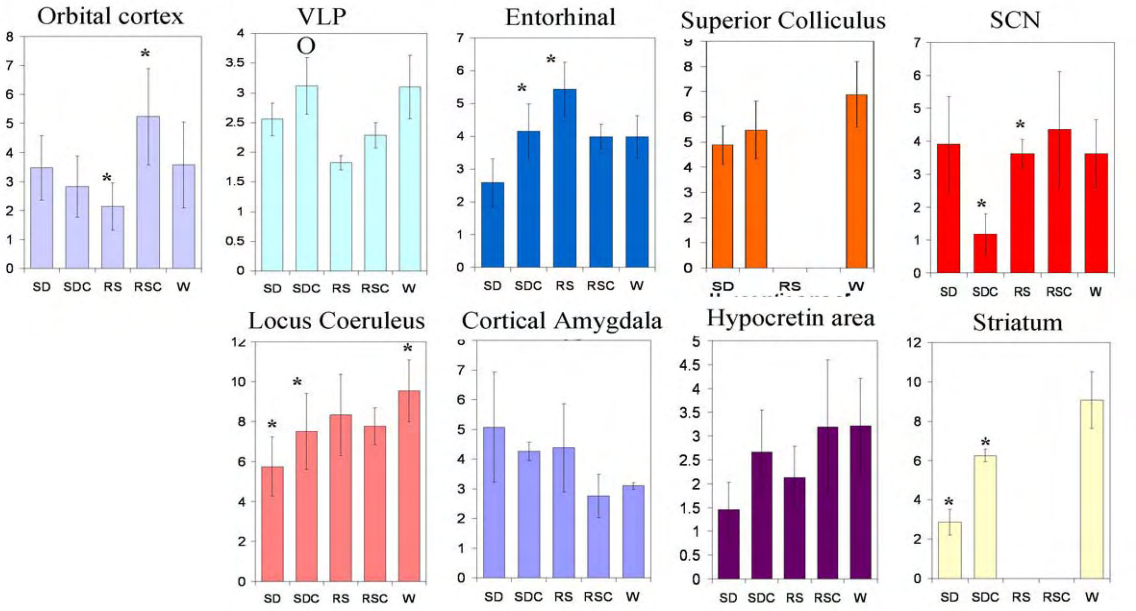


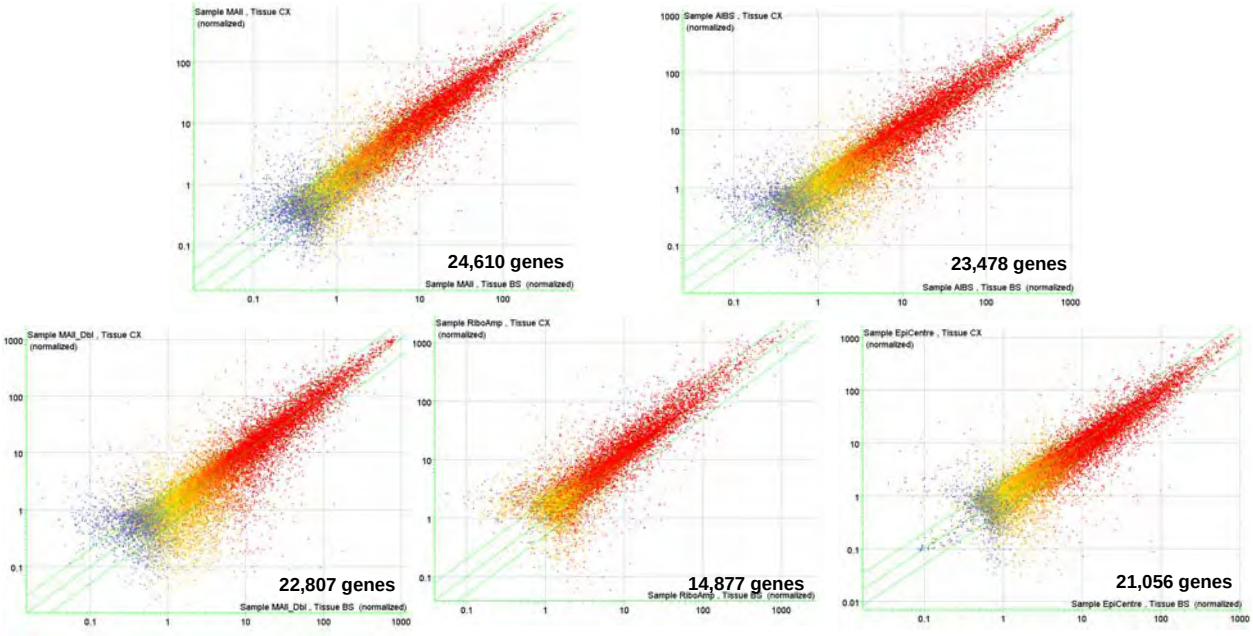
Figure 11. RNA yields vary by sleep condition in a region-specific manner.

Microarrays

High fidelity RNA amplification is essential to obtain reproducible, high quality microarray data. A key element is to have identical input amounts. We chose 5ng total RNA input based on past experience and for pragmatic limitations on the possible yields from structures of interest for this study. Amplifications and microarray processing were outsourced to a service provider, GenUs. A pilot was initially performed with

GenUs to compare different amplification kits using diluted high quality RNA from brainstem and cortex. This RNA was used for both single amplification and double amplification in the Ambion MessageAmp II kit, as well as the Arcturus RiboAmp kit and Epicentre Kit. These different amplification methods gave somewhat varying results. The amplification from the RiboAmp kit did not yield the expected length of amplified message; all other kits gave approximately 500 nt aRNA after double amplification.

Triplicate samples for each amplification method, processed on different days, were applied to GE Codelink Arrays to determine the highest fidelity method. Codelink arrays were selected due to the expertise of GenUs as well as claims of increased sensitivity relative to the Affymetrix high density oligonucleotide array platform. It is anticipated that there will be a loss of sensitivity with two amplification steps compared to a single amplification. A simple means of assessing amplification sensitivity is to examine the percentage of genes called "present" with the different methods. Scatterplots of replicates for each kit are shown in **Figure 12**. The RiboAmp samples had the fewest number of present calls for genes present in at least 2 of 3 samples (65% as many genes as the MessageAmpII samples). However, the other methods had present calls only slightly lower than the single amplification standard. Based on these results, the Ambion MessageAmp II kit with double amplification on 5 ng total RNA starting material was chosen as the method for the remainder of the study.



Genes present (above background) in at least 2 samples are indicated for each amplification protocol
Each spot is the mean of 2 or 3 amplifications for each tissue type
Red=High expression, Yellow=Medium expression, Blue=Low expression

Figure 12. Scatterplots by amplification method. Top row: MessageAmp II single amplification (left), and MessageAmp II double amplification at AIBS (right). Bottom row: MessageAmp II double amplification at GenUS (left), RiboAmp double amplification (middle) and EpiCentre double amplification (right).

The original experimental plan called for four biological replicates of microarrays per behavioral condition for a total of 160 arrays. Five experimental brains per condition were allotted for laser microdissection, and RNA samples from LCM-isolated regions were sent to GenUs for amplification and microarray analysis. **Table 2** indicates the current status of data generation for each structure/behavioral condition.

	Condition				
Region	SD	SDC	RS	RSC	W
HCRT	3	4	3	3	3

LC	3	4	4	4	3
mENT	2	3	4	4	4
ORB	4	4	3	4	4
PMCO	3	3	3	5	3
SCN	4	2	4	4	3
TMN	2	2	3	3	2
VLPO	2	3	3	3	3

Table 2. Summary of completed microarray experiments. SD: sleep deprivation; SDC: sleep deprivation control; RS: recovery sleep; RSC: recovery sleep control; W: waking; HCRT: hypocretin cell containing region of the tuberal hypothalamus; LC: locus coeruleus; mENT: medial entorhinal cortex; ORB: orbitofrontal cortex; PMCO: Posteromedial cortical amygdala; SCN: suprachiasmatic nucleus; TMN: tuberomammillary nucleus; VLPO: ventrolateral preoptic area.

Data reproducibility has been generally quite good. For example, correlation coefficients between 0.92 and 0.97 were obtained for replicates of orbital cortex within each experimental condition (**Figure 13**). As anticipated, there is greater sample to sample variability with heavily amplified samples compared to non-amplified samples, and there is also biological variability associated with individual animals in different behavioral conditions. Despite these sources of variability, the available microarray data indicate clear differences between brain regions and between behavioral conditions, as described below.

There have been several problems associated with generating microarray data. The first problem is that the amplification methodology is near the limit for generating sufficient probe using 5ng starting material. A number of samples have failed amplification and needed to be regenerated. Second, the variance in RNA yield by condition has caused problems obtaining significant quantity of RNA for amplification for several samples. Finally, and quite unfortunately, GE Codelink array production is being discontinued this year, and it has been difficult for GenUs to obtain sufficient arrays to complete the study. It is anticipated that this problem will be resolved, although it is likely that several samples, in particular the VLPO samples that give poor yield, will not be completed as planned. Furthermore, the lack of informatics support for the Codelink platform once the product line is discontinued will make subsequent analysis of the microarray data difficult.

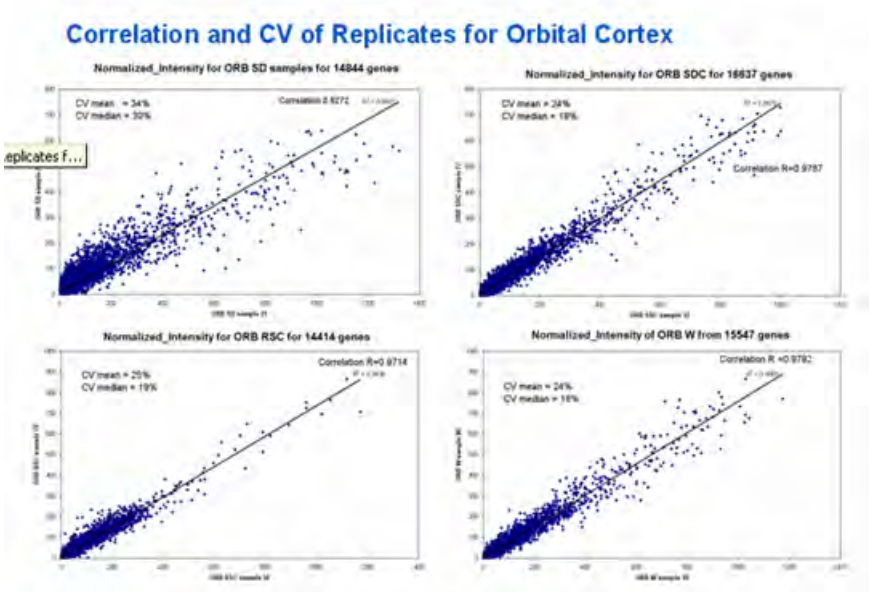


Figure 13. Correlation and CV for orbital cortex replicates.

Aim 2.2: ISH of candidate genes derived from LCM experiments on experimental conditions described in Aim 1.1.

Selection of microarray candidates

As microarray data has become available, genes that are predicted to change as a function of behavioral condition have been selected for microarray analysis. Initial analysis of the data involved filtering to remove outliers from the dataset, and to threshold for expression level. Several strategies have been chosen to select a wide range of candidates. These include searching for genes that show consistent up- or down-regulation by condition within a single structure. Another strategy used was to search for genes with patterns that are similar to genes that are known to change with sleep deprivation (predominantly the immediate early genes assayed earlier). Finally, there is significant sample-to-sample variation that likely reflects a combination of biological variability between individual animals, and amplification variance between samples. Strategies to treat each sample as an independent data point rather than averaging across replicates have proven to be more reliable to find candidates that confirm by ISH. A technical difficulty has been that many samples on the Codelink arrays do not correlate well with other genomic information or correspond with noncoding regions. The ideal candidates genes are those whose sequences code for known transcripts, and, from the perspective of confirmation with ISH, whose sequences overlap with the probes used for the Allen Brain Atlas that are used for confirmation.

Fidelity of the array data was initially assessed by looking for differences between brain regions, pooled across all experimental conditions, that could be confirmed by Allen Brain Atlas data. Qualitatively, these data match the available ISH data. For example, many genes predicted to be enriched in the locus coeruleus show enrichment by ISH. On the other hand, the array data could not predict significant enrichment in the ventrolateral preoptic data, consistent with our inability to find specific markers for this region in the Allen Brain Atlas data. Subsequently, gene lists were generated for genes associated with either sleep deprivation, recovery sleep, or waking. As predicted, immediate early genes are among the more robust genes showing changes between behavioral conditions. For the orbital cortex (n=3), relatively few genes show state-specific enrichment, containing approximately 10 genes. Interestingly, early analysis indicates that the locus coeruleus and hypocretin area samples contain approximately 100-150 genes each associated with one of those sleep conditions. The ten-fold difference in numbers of genes that change with sleep condition may be explained by the fact that the locus coeruleus and hypocretin areas are actually sleep regulatory regions, and thus large changes in gene expression are expected. Furthermore, large fold changes have been observed in the locus coeruleus samples (7-15 fold) with small values for standard deviation.

An example of microarray prediction and ISH confirmation is shown in **Figure 14**. In this case, Dusp14 and Arr3 were predicted to be enriched in recovery sleep (RS) relative to other conditions in posteromedial cortical amygdala (PMCO). ISH data confirm relative enrichment in RS, but also in sleep deprived (SD) state. These data provide important additional information regarding the cell density. In the case of Dusp14, most cells appear to show this upregulation, while Arr3 levels are only increased in a very sparse cell population.

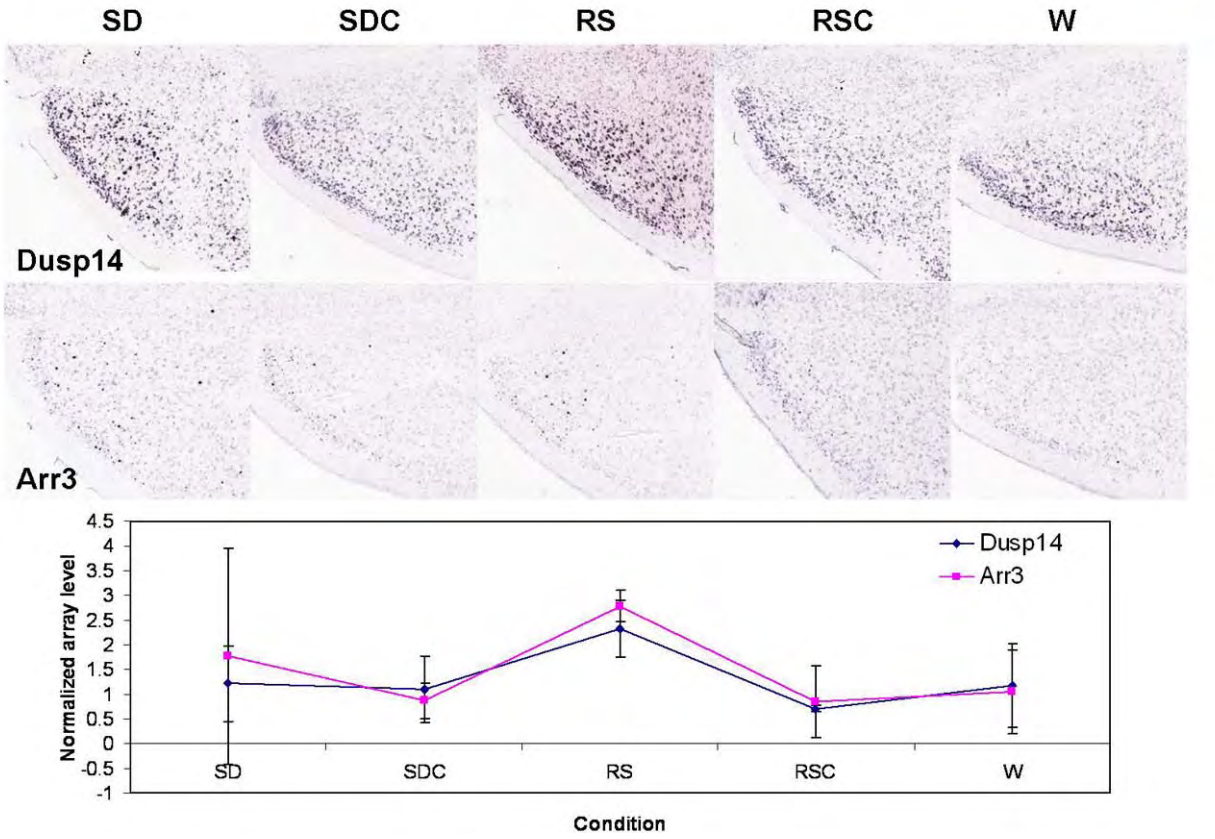


Figure 14. Dusp14 and Arr3 in the posteromedial cortical amygdala (PMCO). A magnified view of the PMCO for each gene across the five behavioral conditions. The array levels for these genes are shown in the graph at the bottom. Array data is n=3, with the exception of condition RSC (n=2).

To date 144 candidate genes have been selected from microarray data for validation (**Appendix II**). This data production is in progress at the present time. It should be noted that a significant proportion of the candidates have not confirmed microarray predictions. The single biggest reason for this discrepancy appears to be a threshold expression level necessary to detect signal by ISH. Not surprisingly, the genes showing the largest fold changes across condition are those with the lowest expression levels. Further filters for minimum expression level have been added to the candidate selection to ensure a higher rate of confirmation. A subset of immediate early genes have now been processed for replicate data sets, and a total of 41 genes will ultimately have replicate data sets generated.

Aim 2.3: Quantification and analysis of ISH results, presentation of data on AIBS publicly accessible web site.

The latter portion of the project (Months 12-18) largely involve data analysis and public presentation of the entire data set. Tools for quantification have been developed as described above, and as the demands of candidate selection and data production ease, a greater focus will be placed on detailed analysis of the larger ISH data set. As mentioned earlier, ISH data from the initial 48 genes derived from the literature is now available on the Allen Brain Atlas web site (**Figure 15**; www.brainatlas.org). The entire data set will be made publicly available through this interface by the end of the project.

From a technical perspective, the project has been very successful to date. Technical development of the sleep deprivation paradigm, laser capture microdissection and array methodologies, and development of a web-based interface for data presentation have proceeded on schedule. Delays in working with the service provider for microarray data generation have pushed back candidate selection and ISH data generation several months from the original plan, but will be completed within the grant period of 18 months.

From a scientific perspective the project is yielding very interesting findings. First, the comprehensive nature of the ISH analysis allowed by our high-throughput methodologies provides an unprecedented analysis of brain regions undergoing selective activation as a function of behavioral state. The careful inclusion of circadian controls has proved to be important, as some of the changes observed are clearly circadian or have to do with exposure of the animals to light at a particular timepoint, for example in visual cortex in light vs. dark. Immediate early gene analysis has indicated several regions not previously focused on as affected by sleep deprivation, including cortical amygdalar regions and the entorhinal cortex, as well as regions that would be expected based upon behavioral observations, such as frontal cortex. Interestingly, different immediate early genes show different patterns of activation, either reflecting different timecourses of activation, or that immediate early gene expression is not strictly a consequence of neuronal activation. Preliminary analysis of the microarray candidate gene ISH indicates a very complex scenario with regional and behavioral state-specific changes in gene expression, often in very discrete cell populations. Data from this study will be presented in two presentations at the Sleep 2007 meeting.

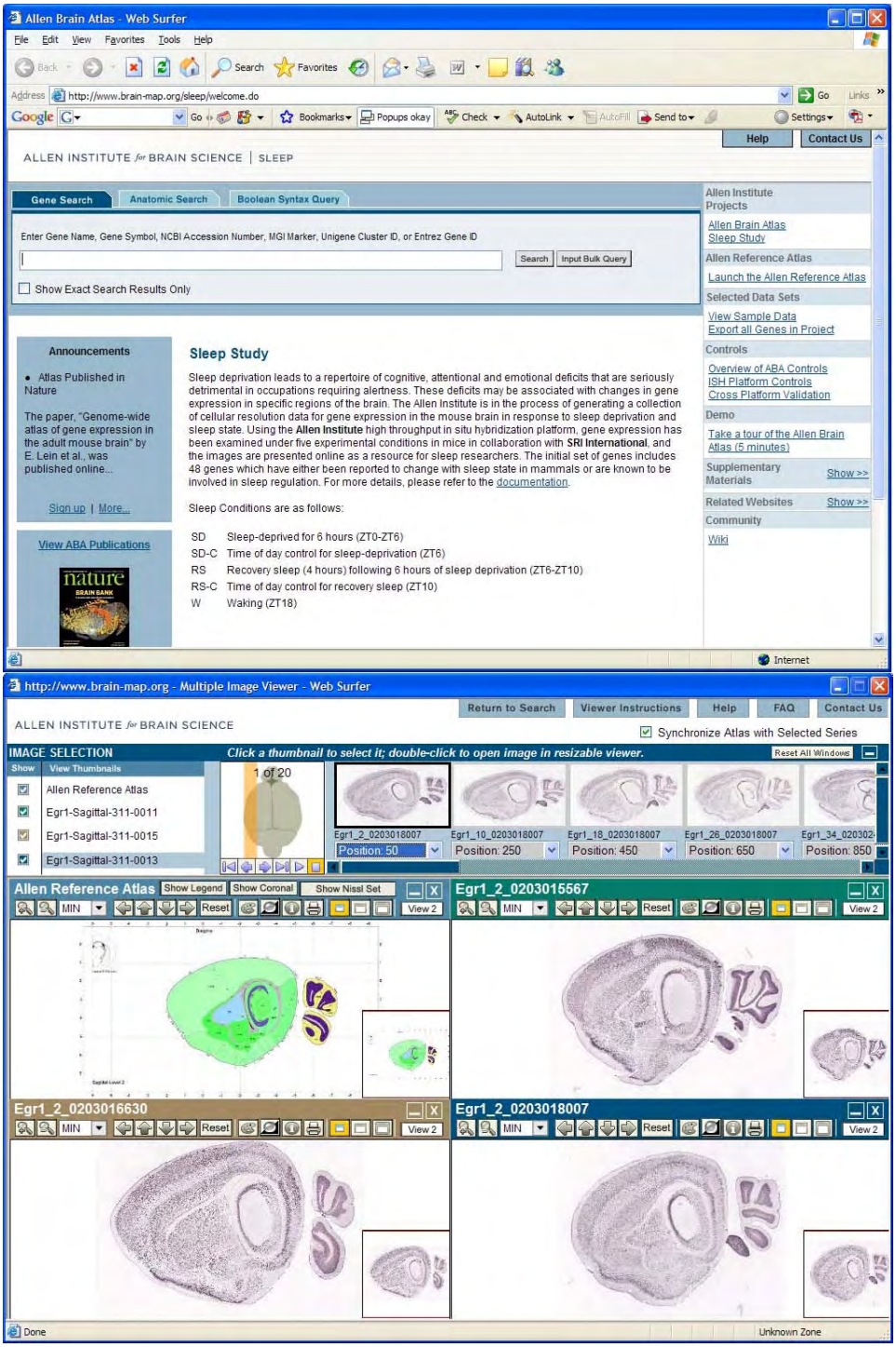


Figure 15. Sleep study web interface for public access to ISH data.

Key Research Accomplishments

Successful implementation of a behavioral paradigm for sleep deprivation and recovery sleep with sufficient controls to dissociate sleep/wake state, sleep deprivation state and circadian time. This paradigm has been validated by EEG/EMG recording as well as by confirmation of gene expression changes associated with these manipulations previously described. Generation of the entire cohort of experimental animals for this study.

Generation of an anatomically comprehensive *in* situ hybridization data set for a set of 48 genes culled from the sleep/sleep deprivation literature, most notably containing immediate early genes.

Generation of a web-based interface allowing open access to the entire raw image data set and comparison between conditions. Public release of data from the initial 48 genes across 5 behavioral conditions.

Identification of brain regions showing most robust gene activation as a function of sleep deprivation and recovery sleep, selection of these regions for microarray analysis.

Development of a robust methodology for isolating small discrete brain regions using laser capture microdissection, preparing high quality RNA from these samples, and amplifying this material for microarray analysis.

Generation of microarray data for ~130 samples spanning 8 brain regions involved in sleep/wake regulation or the effects of sleep deprivation across five behavioral conditions

Selection of candidate genes from microarray analysis (144) and generation of *in situ* hybridization data for ~100 genes to date.

Development of quantitative tools to analyze large-scale image data and quantify changes in expression associated with sleep deprivation.

Reportable Outcomes

Two abstracts have been submitted to the SLEEP 2007 21st Annual Meeting of the APSS meeting:

Mapping of neuronal activity by immediate early gene expression during sleep deprivation and recovery sleep.

Carol L. Thompson[1], Jonathan P. Wisor[2], Dmitry Gerashchenko[2], Shanna Fischer[1], Sayan Pathak[3], Michael Hawrylycz[3], Allan Jones[1], Thomas S. Kilduff[2], and Ed Lein[1].

[1]Neuroscience Department, and [3]Informatics Department, Allen Institute for Brain Science, Seattle, WA; [2]Neurobiology Program, SRI International, Menlo Park, CA

Introduction
Identification of brain regions activated by sleeping, waking and sleep deprivation (SD) may help to understand sleep regulation and deficits associated with SD. Changes in gene expression can reflect altered neuronal activity, particularly rapidly regulated genes like immediate early genes (IEGs). In the current study, high throughput *in situ* hybridization (ISH) was used to map dynamic expression of 48 genes in association with sleeping, waking and sleep-deprived behavioral states.

Methods
Male C57BL/6 mice underwent 6 hr SD (ZT0-ZT6) by cage tapping and introduction of novel objects; a subset were subsequently allowed recovery sleep (RS) for 4 h (ZT6-ZT10). Time of day controls and a spontaneous waking timepoint (ZT18) were collected. ISH was performed on brains for 48 mRNAs including c-*fos, arc, nr4a1, egr3, fosb, junb, egr1, homer1,* and *fosl2* using high throughput colorimetric ISH, and fluorescent co-labeling of c-*fos* with *gad1* mRNA. Images collected on an automated microscopy platform were analyzed by manual and automated image processing.

Results
SD and RS mice were awake >95% during SD; during the hour subsequent to sleep onset after SD, RS mice had greater %NREMS, NREMS delta power and NREMS bout duration than controls. A quarter of the genes exhibited changes in expression correlated with sleep state in regions associated with functional deficits incurred by SD. Prominent patterns included SD-associated gene induction in caudate putamen, cerebellum, visual cortex, and basomedial amygdala, and SD- and wake-associated induction in cortical amygdala and orbital cortex. IEG expression in striasomes varied across conditions. The absence of c-*fos* and *gad1* colabeling in most regions suggests SD preferentially affects excitatory neurons.

Conclusion
Sleep deprivation leads to altered gene expression in brain regions associated with cognitive, emotional, and memory deficits. Similar activation patterns are observed across immediate early genes in response to sleep deprivation and/or recovery sleep, with a few gene-specific differences.

Support
Supported by the Department of Defense USAMRAA award W81XWH-06-1-0131 and NIH 5RO1HL59658.

Gene expression patterns in the locus coeruleus during sleep/wakefulness: microarray analysis

Dmitry Gerashchenko[1], Carol Thompson[2], Jonathan P. Wisor[1], Lydia Ng[2],
Kimberly Smith[2], Allan Jones[2], Thomas Kilduff[1], Ed Lein[2]

[1] Neurobiology Program, SRI International, Menlo Park, CA; [2] Allen Institute for Brain Science, Seattle, WA

Introduction

Noradrenergic locus coeruleus (LC) neurons play an important role in arousal regulation. Firing of noradrenergic LC neurons is high during wakefulness, progressively decreases during slow wave sleep, and becomes practically silent during paradoxical sleep. It is expected that state-associated changes in LC activity are associated with corresponding changes in gene expression in LC neurons. Identification of gene expression patterns in LC neurons could help to understand mechanisms of sleep/wake regulation. In the current study, we studied gene expression profile in the LC by identifying the genes whose expression is enriched in the LC region and by doing microarray analysis of gene expression in the LC during different sleep/wake states.

Methods

The Allen Brain Atlas was used to find genes that are expressed stronger in the LC than in the surrounding brain areas. Digitized intensity of *in situ* hybridization signal in the LC was divided by the intensity of signal in the pontine area, and the genes with highest ratios were selected. For microarray analysis, LC samples were collected by using the laser-capture microdissection in mice in the following 5 groups: 1) 6 hours of sleep deprivation (ZT0-ZT6), 2) control for sleep deprivation (ZT6), 3) 6 hours of sleep deprivation followed by recovery sleep (ZT10), 4) control for recovery sleep (ZT10), 5) spontaneous waking (ZT18). cDNA synthesis, cRNA amplification and hybridization on chips were done according to standard CodeLink protocols. Data were analyzed using ANOVA.

Results

Approximately 2000 genes that are enriched in the LC have been selected from the Allen Brain Atlas based on the highest expression ratios of LC /brainstem *in situ* hybridization signal. According to gene ontology classification, a lager part of these genes are involved in the cellular physiological processes. The microarray results showed that the largest number of genes that strongly changed their expression was observed in the sleep deprivation group. Most of differentially expressed genes that were also present on the list of genes enriched in the LC belonged to either sleep deprivation group (*dexi, epdr2, cul4a, prnp, neurl*) or spontaneous waking group (*prnp, lrrn6c, nrg1, pgrmc1, neurl*).

Conclusion

Gene expression results of the present experiment are consistent with an important role of the LC in arousal regulation.

Conclusion

The results of the current study demonstrate that the behavioral deficits associated with sleep deprivation are mirrored by changes in gene expression. As predicted, structures involved in higher cognitive functions (neocortex), memory (hippocampal formation) and emotional behavior (amygdala) all show altered gene expression in response to sleep deprivation. However, these changes in gene expression are local, occurring in very discrete brain regions and nuclei, rather than global, and often occur in small cell populations within these regions. Furthermore, the function of most of these genes, as they relate to sleep biology, is completely unknown. Robust changes are also found in regions associated with sleep/wake regulation. Understanding both the genes undergoing dynamic regulation as well as the structures and cell classes in which they are expressed is critical to generating a mechanistic view of how sleep deprivation leads to impaired behavior, and how subsequent sleep reverses the process. The data generated from the current project will serve a number of purposes. First, it will provide a comprehensive data set for the genes analyzed, allowing the entire sleep research community to benefit by studying novel identified genes and their role in sleep regulation and dysregulation. Second, these data serve as a baseline with which to compare pharmacological or genetic manipulations that affect sleep regulation or alleviate the consequences of sleep deprivation. Further analysis of these data will hopefully help to identify specific biochemical pathways affected by sleep deprivation as a means to generate new targets for pharmacological intervention.

References

1. D. Leger, *Sleep* **17**, 84 (Feb, 1994).
2. H. P. Van Dongen, G. Maislin, J. M. Mullington, D. F. Dinges, *Sleep* **26**, 117 (Mar 15, 2003).
3. A. Oliverio, *Waking Sleeping* **4**, 155 (Apr-Jun, 1980).
4. S. C. Veasey *et al.*, *Sleep* **23**, 1025 (Dec 15, 2000).
5. A. Terao, M. A. Greco, R. W. Davis, H. C. Heller, T. S. Kilduff, *Neuroscience* **120**, 1115 (2003).
6. P. Franken, D. Chollet, M. Tafti, *J Neurosci* **21**, 2610 (Apr 15, 2001).
7. P. Franken, A. Malafosse, M. Tafti, *Sleep* **22**, 155 (Mar 15, 1999).
8. C. Cirelli, C. M. Gutierrez, G. Tononi, *Neuron* **41**, 35 (Jan 8, 2004).
9. C. Cirelli, *J Appl Physiol* **92**, 394 (Jan, 2002).
10. A. Terao *et al.*, *Sleep* **23**, 867 (Nov 1, 2000).
11. A. Terao *et al.*, *Neuroscience* **116**, 187 (2003).
12. A. A. Borbely, I. Tobler, M. Hanagasioglu, *Behav Brain Res* **14**, 171 (Dec, 1984).
13. T. Deboer, P. Franken, I. Tobler, *J Comp Physiol [A]* **174**, 145 (Feb, 1994).
14. R. Huber, T. Deboer, I. Tobler, *Brain Res* **857**, 8 (Feb 28, 2000).
15. I. Tobler, K. Jaggi, *J Comp Physiol [A]* **161**, 449 (Aug, 1987).

Appendices

Appendix I. 48 candidate genes derived from sleep/circadian literature.

Gene Name	Gene Symbol	Chromo-some	NCBI Accession	MGI Marker Accession Id
adenosine A1 receptor	Adora1	1	NM_009629	MGI:99401
activity regulated cytoskeletal-associated protein	Arc	15	NM_018790	MGI:88067
aryl hydrocarbon receptor nuclear translocator-like	Arntl	7	NM_007489	MGI:1096381
brain derived neurotrophic factor	Bdnf	2	NM_007540	MGI:88145
cholecystokinin	Cck	9	NM_031161	MGI:88297
choline acetyltransferase	Chat	14	NM_009891	MGI:88392
cholinergic receptor, nicotinic, beta polypeptide 2 (neuronal)	Chrnb2	3	NM_009602	MGI:87891
cardiotrophin-like cytokine factor 1	Clcf1	19	NM_019952	MGI:1930088
cannabinoid receptor 1 (brain)	Cnr1	4	NM_007726	MGI:104615
cortistatin	Cort	4	NM_007745	MGI:109538
corticotropin releasing hormone	Crh	3	NM_205769	MGI:88496
casein kinase 1, alpha 1	Csnk1a1	18	NM_146087	MGI:1934950
casein kinase 1, epsilon	Csnk1e	15	NM_013767	MGI:1351660
dopamine beta hydroxylase	Dbh	2	NM_138942	MGI:94864
early growth response 1	Egr1	18	NM_007913	MGI:95295
early growth response 3	Egr3	14	NM_018781	MGI:1306780
FK506 binding protein 1a	Fkbp1a	2	NM_008019	MGI:95541
FBJ osteosarcoma oncogene	Fos	12	NM_010234	MGI:95574
FBJ osteosarcoma oncogene B	Fosb	7	NM_008036	MGI:95575
fos-like antigen 2	Fosl2	5	NM_008037	MGI:102858
glutamic acid decarboxylase 1	Gad1	2	NM_008077	MGI:95632
galanin	Gal	19	NM_010253	MGI:95637
glial fibrillary acidic protein	Gfap	11	NM_010277	MGI:95697
growth hormone releasing hormone	Ghrh	2	NM_010285	MGI:95709
histidine decarboxylase	Hdc	2	NM_008230	MGI:96062
homer homolog 1 (Drosophila)	Homer1	13	NM_011982	MGI:1347345
heat shock protein 90kDa beta (Grp94), member 1	Hsp90b1	10	NM_011631	MGI:98817
heat shock 70kD protein 5 (glucose-regulated protein)	Hspa5	2	NM_022310	MGI:95835
Jun-B oncogene	Junb	8	NM_008416	
neuronal PAS domain protein 2	Npas2	1	NM_008719	MGI:109232
neuronal pentraxin 2	Nptx2	5	NM_016789	MGI:1858209
nuclear receptor subfamily 4, group A, member 1	Nr4a1	15	NM_010444	MGI:1352454
nuclear receptor subfamily 4, group A, member 3	Nr4a3	4	Mm.96708	MGI:1352457
proprotein convertase subtilisin/kexin type 2	Pcsk2	2	NM_008792	MGI:97512
protein disulfide isomerase associated 4	Pdia4	6	NM_009787	MGI:104864
prodynorphin	Pdyn	2	NM_018863	MGI:97535
period homolog 1 (Drosophila)	Per1	11	NM_011065	MGI:1098283
period homolog 2 (Drosophila)	Per2	1	NM_011066	MGI:1195265
period homolog 3 (Drosophila)	Per3	4	NM_011067	MGI:1277134
pro-melanin-concentrating hormone	Pmch	10	NM_029971	MGI:97629
protein phosphatase 1, regulatory (inhibitor)	Ppp1r3c	19	NM_016854	MGI:1858229

subunit 3C				
prokineticin 2	Prok2	6	NM_015768	MGI:1354178
retinoic acid receptor, beta	Rarb	14	NM_011243	MGI:97857
RAS, dexamethasone-induced 1	Rasd1	11	NM_009026	MGI:1270848
tyrosine hydroxylase	Th	7	NM_009377	MGI:98735
tryptophan hydroxylase 2	Tph2	10	NM_173391	MGI:2651811
thyrotropin releasing hormone	Trh	6	NM_009426	MGI:98823
vasoactive intestinal polypeptide	Vip	10	NM_011702	MGI:98933

Appendix II. Genes selected from microarray data for ISH confirmation

Gene Symbol	Gene Name	Condition	Predicted Region
Fut8	Fucosyltransferase 8	RS associated	MENT
Zbtb7c	(B230208J24Rik)	SD associated	TMN
Tmem166	(BC014699)	RS associated	HCRT
0610037L13Rik	0610037L13Rik	SD and RS associated	ORB
1300007B12Rik	1300007B12Rik	RS associated	LC
4930418G15Rik	4930418G15Rik	RS associated	MENT
Tarsl2	A530046H20Rik	W associated	SCN
Acss1	Acetyl-Coenzyme A synthetase 2 (AMP forming)-like	W associated	LC
Ap1m2	Adaptor protein complex AP-1, mu 2 subunit	SD associated	Hcrt
Add3	Adducin 3 (gamma)	W associated	PMCO
Ak2	Adenylate kinase 2	SD associated	MENT
Aard	Alanine and arginine rich domain containing	SD associated	ORB
Agtr1a	Angiotensin receptor 1	RS, SD, and W associated	LC
Asb8	Ankyrin repeat and SCOS box-containing protein 8	W associated	PMCO
Ankrd9	Ankyrin repeat domain 9	RS associated	MENT
Ankrd17	ankyrin repeat domain protein 17 isoform	W associated	PMCO
Mki67	Antigen identified by monoclonal antibody Ki 67	SD associated	MENT
Kin	Antigenic determinant of rec-A protein	SD, W associated	ORB
Actr3	ARP3 actin-related protein 3 homolog	W associated	ORB
Arr3	Arrestin 3, retinal	RS associated	PMCO
Atp5a1	ATP synthase, H+ transporting, mitochondrial F1 complex, alpha subunit, isoform 1	SD associated	MENT
Abca5	ATP-binding cassette, sub-family A (ABC1), member 5	SD associated	PMCO
Bace2	Beta-site APP-cleaving enzyme 2	W associated	MENT
Pygb	Brain glycogen phosphorylase	SD down	LC
Baz1b	Bromodomain adjacent to zinc finger domain, 1 B	RS associated	LC
Bub3	Budding uninhibited by benzimidazoles 3	RS associated	PMCO
Cdh17	Cadherin 17	W associated	LC
Car8	Carbonic anhydrase 8	RS down	TMN
Comt	Catechol-O-methyltransferase	SD associated	ORB MENT PMCO
Ccrn4l	CCR4 carbon catabolite repression 4-like (S. cerevisiae)	RS associated	LC
Cnot6l	CCR4-NOT transcription complex, subunit 6-like	RS associated	ORB
Cul4a	Cullin 4a	SD down	LC
Cdkn1a	Cyclin-dependent kinase inhibitor 1A	SD associated	PMCO, Hcrt, LC, SCN
Cpeb1	Cytoplasmic polyadenylation element binding protein 1	SD associated	MENT
Cova1	Cytosolic ovarian carcinoma antigen 1	RSC and SDC associated	SCN
D930016N04Rik	D930016N04Rik	RS associated	LC
Dctd	DCMP deaminase	W associated	VLPO
Dbi	Diazepam binding inhibitor	SD and RS associated	PMCO

Dscr3	Down syndrome critical region gene 3	RS associated	HCRT
Dusp4	Dual specificity phosphatase 4	SD, RS, and W associated	MENT
Dusp14	Dual specificty phosphatase 14	RS associated	PMCO
Ebf3	Early B-Cell factor 3	RS associated	Hcrt
Egr2	Early growth response 2	SD and W associated	ORB
Egln2	EGL nine homolog 2	SD associated	MENT PMCO ORB
Epdr2	Epenydmin 2 (AU040950)	SD associated	LC
Ebag9	Estrogen receptor-binding fragment-associated gene 9	RS associated	PMCO
Fbxo28	F-box protein 28	SD associated	ORB
Fgl1	Fibrinogen-like protein 1	SD, RS, and W associated	LC
Mical3	Flavoprotein oxidoreductase MICAL3	W associated	LC
Fhod1	Formin homology 2 domain containing 1	RS associated	MENT
Gpr75	G protein-coupled receptor 75	SD associated	ORB
Gabra2	Gamma-aminobutyric acid (GABA-A) receptor	SD associated	MENT
Gabrg2	Gamma-aminobutyric acid (GABA-A) receptor, subunit gamma 2	RS associated	MENT
Pdia3	Glucose regulated protein (Grp58)	SD, RS, and W associated	ORB
Gfpt2	Glutamine fructose-6-phosphate transaminase 2	W associated	HCRT LC
Gldc	Glycine decarboxylase	W associated	LC
Gramd2	GRAM domain containing 2 (4933407I18Rik)	W down	VLPO
Gas1	Growth arrest specific 1	RS associated	Hcrt
Gda	Guanine deaminase	SD associated	SCN
Hspb1	Heat shock protein 1	SD and W associated	MENT
Hnrpf	Heterogenous nuclear ribonucleoprotein F	SD associated	SCN
H47	Histocompatibility 47	SD associated	MENT
Ppa2	inorganic pyrophosphatase 2 (1110013G13Rik)	RS associated	Hcrt
Kras	Kirsten rat sarcoma oncogene 2, expressed (Kras2)	RS associated	ORB
Me2	Malic enzyme 2, NAD(+)-dependent, mitochondrial	SD associated	MENT
Mettl6	methyltransferase like 6 (1600013P15Rik)	RS associated	LC
Mybpc2	Myosin binding protein C, fast-type	W associated	SCN
Nedd1	Neural precursor cell expressed, developmentall down-regulated gene 1	W associated	LC
Nedd4	Neural precursor cell expressed, developmentally downregulated gene	SD associated	HCRT LC SCN
Npdc1	Neural proliferation, differentiation and control gene 1	RS associated	ORB
Nrgn	Neurogranin	RS down	LC
Nfatc1	Nuclear factor of activated T-cells, cytoplasmic, calcineurin-dept 1	RS associated	TMN
Nvl	Nuclear VCP-like	SD, RS, and W associated	ALL
Onecut2	One cut domain, family member 2	RS and SD associated	MENT
Ogdh	Oxoglutarate dehydrogenase (lipoamide)	SD associated	LC
Papd1	PAP associated domain containing 1	SD associated	LC

Pwp1	Peroidic tryptophan protein 1 homolog (2310058A11Rik)	SD associated	PMCO
Pgam2	Phosphoglycerate mutase 2	SD associated	PMCO
Phlda1	Pleckstrin homology-like domain, family A, member 1	SD associated	LC
Kcnab2	Potassium voltage-gated channel, shaker-related subfamily, beta member 2	W associated	ORB MENT PMCO
Pgrmc1	Progesterone receptor membrane component 1	RS associated	PMCO
Prokr2	Prokineticin receptor 2, G protein-coupled receptor 73like 1 (Gpr73l1)	W associated	MENT
Pa2g4	Proliferation-associated 2G4	RS associated	HCRT
Pfc	Properdin	SD and W associated	ORB
Htra1	Protease, serine, 11 (Igf binding) (Prss11)	SD associated	LC
Psme1	Proteasome (prosome, macropain) 28 subunit, alpha	RS associated	PMCO
Psmb7	Proteasome (prosome, macropain) subunit, beta type 7	RS associated	PMCO
Psmd2	Proteasome 26S non-ATPase subunit 2	RS associated	HCRT
Pias2	protein inhibitor of activated STAT 2 (Miz1)	RS associated	SCN
Pkia	Protein kinase inhibitor, alpha	SD associated	PMCO
Pxmp4	Proxisomal membrane protein 4	RS associated	MENT
P2ry14	Purinergic receptor P2Y, G-protein coupled, 14	SD associated	TMN
Ranbp2	RAN binding protein 2	RS and W associated	PMCO
Rapgef3	Rap guanine nucleotide exchange factor (GEF) 3	SD associated	ORB
Rgs20	Regulator of G-protein signaling 20	W associated	MENT
Rgs5	Regulator of G-protein signaling 5	SDC and RSC associated	ORB
Rfc5	Replication factor C	RS associated	VLPO
Arhgap18	Rho GTPase activating protein 18	W associated	LC
Arhgap8	Rho GTPase activating protein 18	W associated	LC
Rps6ka2	Ribosomal protein S6 kinase, polypeptide 2	SD associated	MENT PMCO ORB
Rnmt	RNA (guanine-7-) methyltransferase	RS associated	Hcrt
Sav1	Salvador homolog 1 (Drosophila)	W associated	MENT
Sec61a2	Sec61, alpha subunit 2 (S. cerevisiae)	RS associated	MENT
Serpinb1b	Serine (or cysteine) proteinase inhibitor, clade B, member 1b	SD associated	LC
Sgk	Serum/glucocorticoid regulated kinase	SD associated	LC
Neu4	Sialidase 4	SD associated	LC
Scube2	Signal peptide, CUB domain, EGF-like 2	RS and W associated	HCRT
Snf1lk	SNF1-like kinase	SD associated	MENT PMCO ORB
Slc1a2	solute carrier family 1 (glial high affinity glutamate transporter), member 2	SD associated	ORB
Slc5a7	Solute carrier family 5 (choline transporter), member 7	W associated	SCN
Slc6a17	Solute carrier family 6 (neurotransmitter transporter), member 17	SD associated	LC
Snx3	Sorting nexin 3	SD associated	MENT
Stac	Src homology three (SH3) and cysteine rich domain	W associated	SCN
Stab1	stabilin 1	RS associated	HCRT SCN

Tgtp	T-cell specific GTPase	W associated	MENT
Akt3	Thymoma viral proto-oncogene 3	RS associated	SCN
BB328881	Transcribed locus, highly similar to NP_112269.1 adenylyl cyclase	SD associated	LC HCRT
Tcerg1l	Transcription elongation regulator 1-like	SD and W associated	LC
Tcf12	Transcription factor 12	W associated	LC
Tsc22d1	Transforming growth factor beta 1 induced transcript 4 isoform 2 (Tgfb1i4)	RSC and SDC associated	ORB
Tmco1	Transmembrane and coiled-coil domains 1	SD, RS, and W associated	MENT
Tnrc4	Trinucleotide repeat containing 4	SD associated	MENT
Vars2	Valyl-tRNA synthetase 2	RS associated	SCN
Vegfa	vascular endothelial growth factor A	W associated	PMCO
Vgll4	vestigial like 4	SD associated	MENT
Vwf	Von Willebrand factor homolog	SD associated	LC
Zfp235	Zinc finger protein 235	W associated	LC
Zdhhc2	Zinc finger, DHHC domain containing 2	SD and RS associated	HCRT
2610019N13Rik		SD associated	LC
4933422E07Rik		RS associated	MENT
2610002M06Rik		RS associated	LC
D2Ertd391e		RS associated	VLPO
1700010I14Rik		RS associated	VLPO
Ctbs		RS associated	Hcrt
E130009J12Rik		RSC and SDC associated	ORB
LOC215714		RSC and SDC associated	VLPO
4631426J05Rik		RS associated	PMCO
A330104H05Rik		RS down	ORB
2810022L02Rik		RS associated	HCRT LC SCN
2500003M10Rik		SD and RS associated	ORB MENT PMCO
A930038C07Rik		SDC and W up	TMN
1700021L22Rik		RS associated	LC
4931407K02Rik		SD associated	MENT PMCO ORB
Tmem23		RS associated	PMCO